I-Vector Based Depression Level Estimation Technique

Thesis submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering

in

Information Security

Submitted by

Barkha Rani

(801433006)

Under the Supervision of

Dr. Jhilik Bhattacharya

Assistant Professor

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

PATIALA – 147004

June 2016

CERTIFICATE

I hereby certify that the work which is being presented in the thesis entitled, "*I-Vector Based Depression Level Estimation Technique*", in partial fulfillment of the requirements for the award of degree of Master of Engineering in Information Security in Computer Science and Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of Dr. Jhilik Bhattacharya and refers other researcher's work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

(Barkha Rani)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

(Dr. Jhilik Bhattacharya)

Assistant Professor, Computer Science and Engineering Department

Thapar University

Countersigned by

(Dr. Deepak Garg)

Head,CSED

Thapar University

Patiala

(Dr. S.S. Bhatia)

Dean (Academic Affairs)

Thapar University

Patiala

ACKNOWLEDGEMENT

Words are often too less to reveals one's deep regards. An understanding of the work like this is never the outcome of the efforts of a single person. I take this opportunity to express my profound sense of gratitude and respect to all those who helped me through the duration of thesis.

This work would not have been possible without the encouragement and able guidance of my supervisor, Dr. Jhilik Bhattacharya, Assistant Professor, CSED, Thapar University, Patiala. Most of the novel ideas and solutions found in this thesis are the result of our numerous stimulating discussions. Their feedback and editorial comments were also invaluable for the writing of this thesis.

I would also like to thank all the faculty members of the department and my friends who directly or indirectly helped me in completion of my thesis. No words of thanks are enough for my dear parents and brother whose support and care makes me stay on earth. Thanks to be with me.

Above all, I would like to thank the Almighty God for his blessings and for driving me with faith, hope and courage in thinnest of the times.

Place: Thapar University, Patiala

Barkha Rani

ABSTRACT

Depression is considered as a psychosomatic state associated with the soft biometric features. People suffering from depression always behave abnormal. Depression is a clinically proven disorder that can overwhelm a person and his ability to perform even a simple task. Soft biometric provides important information about a person without being enough for their verification because they lack uniqueness. This statement comprises of features which are associated with the psychosomatic state of a person such as feelings, sentiments or brain related disorders like depression. In this dissertation we have estimated the depression level of each speech signal using I-Vector technique. In our proposed approach, first of all we have removed silence from the speech signal then we have extracted features from audio signals using I-Vector, after that split overlapping function is applied to evaluate overlapped audio beats. In the end we have evaluated depression using relationship matrix. We have estimated the depression level of each speaker. This technique has better performance as compared with existing techniques. The overall result has shown that the I-Vector technique has good accuracy to detect depression in audios.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Depression is clinically proven disorder that can overwhelm a person and his ability to perform even a single task. Soft biometric provides important information about a person without being enough for their verification because they lack uniqueness. This statement comprises of features which are associated with the psychosomatic state of a person such as feelings, sentiments or brain related disorders like depression. Depression is a severe mental health problem. People suffering from depression have low feeling, depressed mood, poor spark and poor concentration and always behave abnormally. It is increasing day by day due to load and these issues become persistent. It also destructs human's ability to carry out his or her daily task. It can lead to self-immolation, a disastrous death related with losing thousands of lives every year. It is a serious medical problem and affects the person's behavior and thoughts. Everybody feels depressed from time to time in their life. If any person feels this type of sadness for more than two weeks than that person is suffering from clinical depression. Audio signals also help therapist to measure depression.

Figure 1.1 shows the positron emission tomography (PET) scan of the brain. The depressed part of the brain of a person is showing low activity in the brain as compared to the person who is not depressed. Depression usually starts at young age. It affects female more usually than male, and jobless person are also at high risk. It requires long time treatment.

Figure 1.2 depicts the PET scans of brain. In the diagram it is showing the brain energy consumption rises and falls with emotional switches.

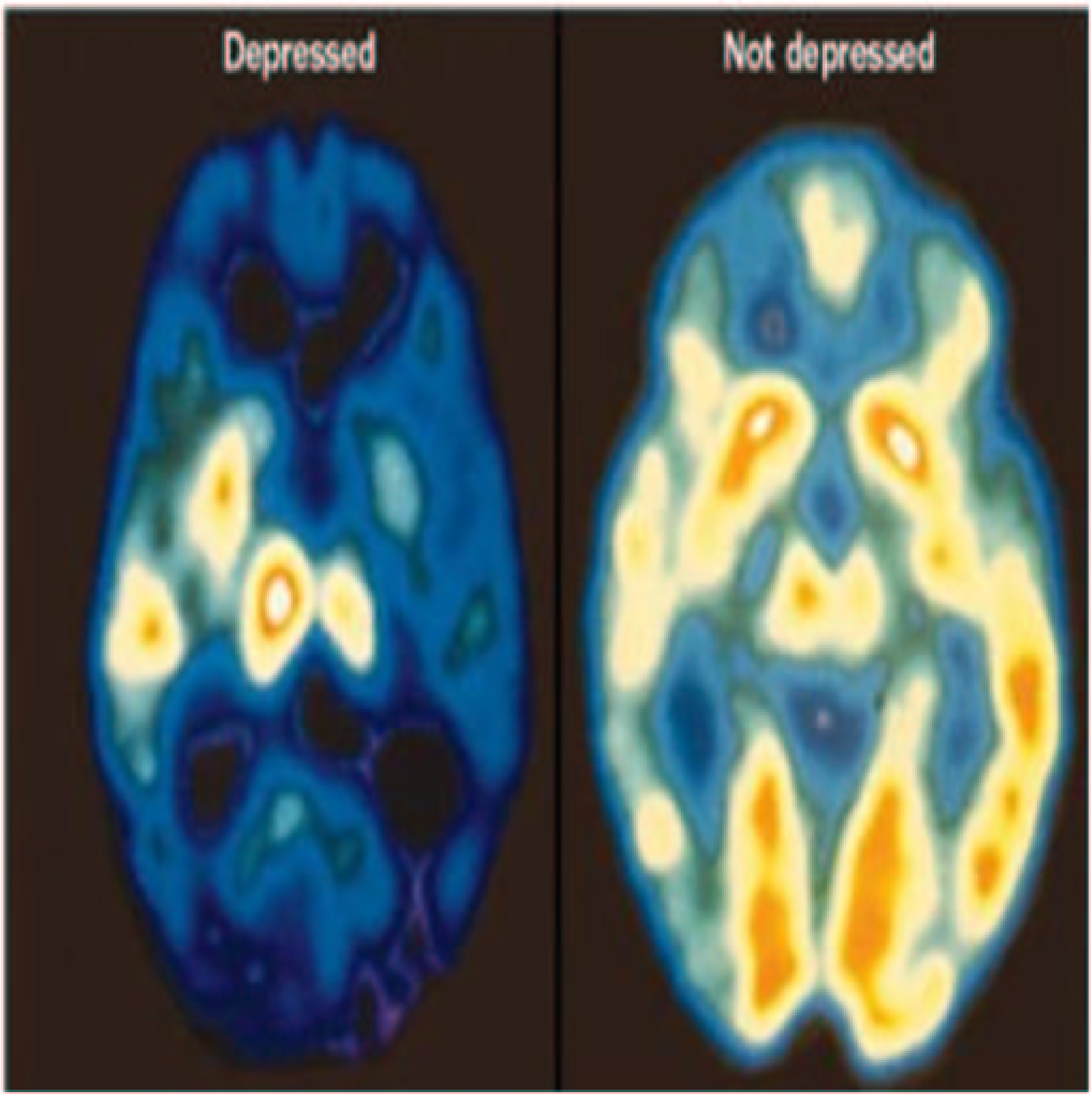

Figure 1.1: Diagram showing PET scans of Brain

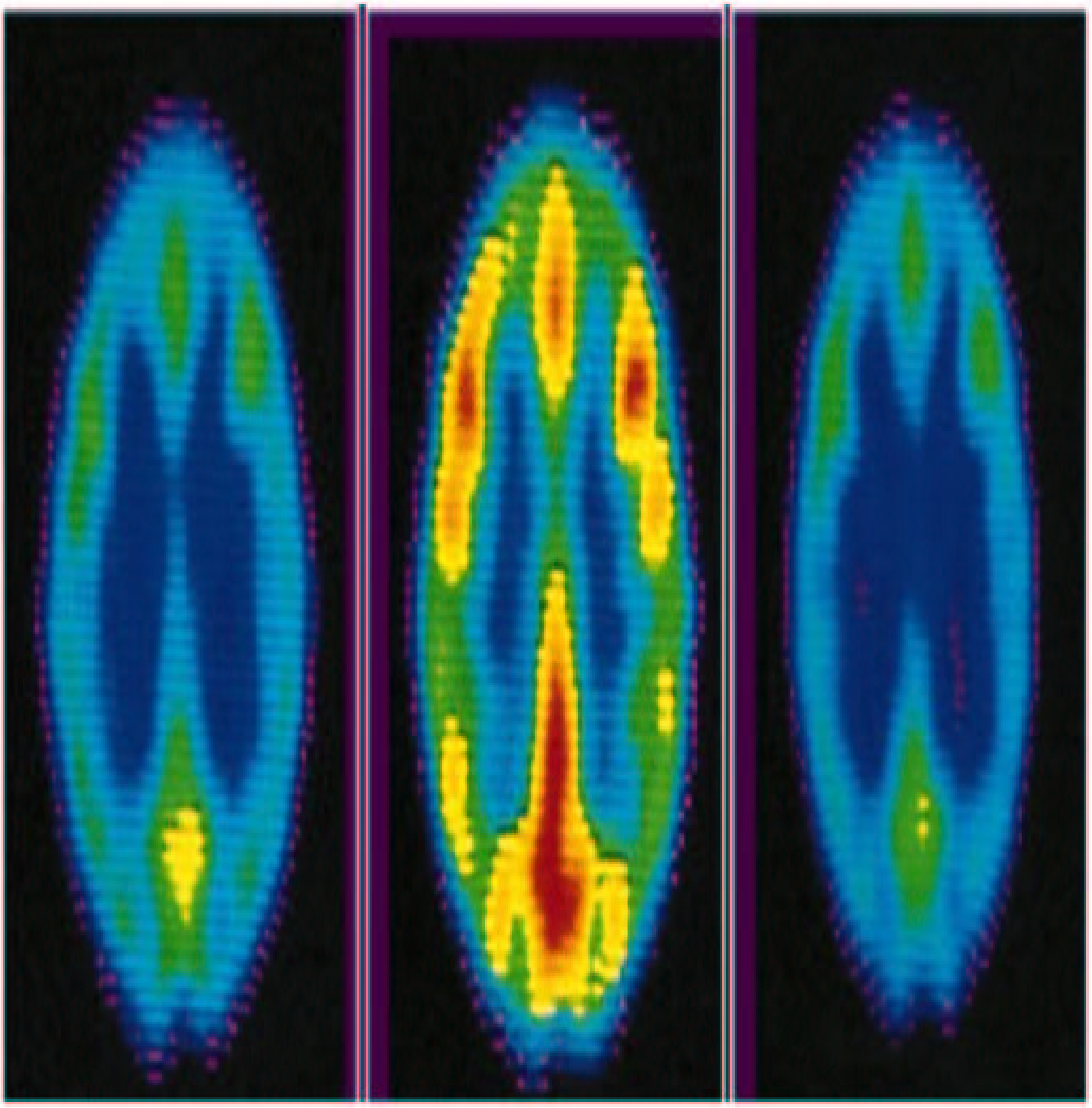

Figure 1.2: Diagram showing Bipolar Mood Disorder Scans of Brain

1.1 Symptoms of Depression

The symptoms of depression differ from one person to another person, yet there are few indications and characteristics. According to "National Institute of Mental Health" following are the symptoms of depression which may include:

- Bad mood, worried, regretful.
- Feelings of hopelessness, disappointment.
- Low energy, tiredness.
- Low concentration and difficulty in making decisions.
- Uneasiness, irritability.
- Thoughts of dying, self-immolation.
- Insomnia, loss of interest, self-destruction.

1.2 Types of Depression

There are various forms and shapes of depression. Knowing different categories of depression can deal with different consequences of depression.

1.2.1 Persistent Depressive Disorder

This type of disorder is also called as dysthymia. People suffering from this type of disorder have depressed mood which lasts for more than two years. In this case symptoms are not serious as like in major depression. This type of symptom must last for two years to be considered persistent depressive disorder.

1.2.2 Perinatal Depression

This type of depression can affect women during pregnancy or generally experience after giving birth to a child. It affects total of 10 to 20 per cent of women during their pregnancy.

Indication of perinatal depression includes shouting, sleep problems, stress, impatience and heart palpitations.

1.2.3 Psychotic Depression

This type of depression occurs when an individual has critical depression and along with loss of touch with reality. The people suffering from major depression also have psychotic depression. Symptoms related with psychotic depression includes: unhappiness, hopelessness, self-accusation and short temperedness.

1.2.4 Seasonal Affective Disorder

This type of depression is related with the changes in seasons. It comes and goes in seasonal pattern. SAD is sometimes known as winter depression because the symptoms are more apparent and tend to be more severe during the winter. Symptoms related with seasonal affective disorder includes: nausea, tendency to oversleep and over eat, persistent low mood and worthlessness.

1.2.5 Bipolar Disorder

This type of disorder is different from depression and also known as manic-depressive illness. It is still included in the list of depression because the person suffering from bipolar disorders experiences episodes of extremely low moods which meets the criteria for major depression. The age at which symptoms begin is 25.

1.2.6 Major Depression

A person who is suffering from this type of depression must either have depressed mood or loss of interest or pleasure in daily activities consistently for at least a 2 week period. Symptoms related with major depression includes: inability to sleep, loss of energy, poor concentration and recurrent thoughts of death.

1.3 Causes of Depression

Unbalancing of chemicals in the body is not only the reason for causing depression. There are different other reasons that can cause depression and it cannot be cured with medicines only. Depression can be caused by following factors:

- **Serious Medical Illness:** Due to stress and worry of coping with a serious illness can lead to depression.
- **Drug and Alcohol Use:** Depression can also occur through excessive intake of drugs and alcohol.
- **Personality:** Few people have possibly more chances of depression due to their nature, especially if the people have capacity to panic or have low self confidence or very delicate to personal disapproval or have negative thinking.
- The problem of depression can also occur through persistent money problem.
- **Loneliness:** Depression can also occur through isolation. The risk of depression can be increased when any person cut off from their families.
- **Family History:** If someone in your family has ever suffered from depression, like your parents or siblings, then there will be more chances that you too will suffer the same way.
- Depression can also occur through serious conflicts.
- **Environmental Factors:** Depression can also occur through environmental disorders such as social isolation, death of loved ones, demanding work or stressful workplace, financial difficulties or job loss.
- **Biological Factors:** Neurotransmitter is a naturally occurring chemical in the body that causes imbalance in the brain and spinal cord. Serotonin and norepinephrine are two neurotransmitters in the brain that appear to be involved in the symptoms of depression. Some traits of depression are same for men and women but women tend to experience more traits then men.

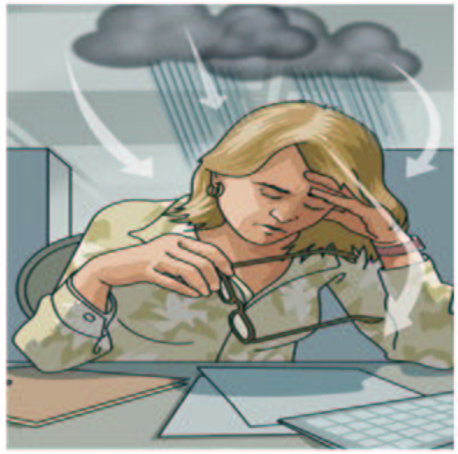

Figure 1.3: Diagram showing women suffering from depression due to excessive workload

1.4 Depression Estimation Strategies

Depression has an important contribution to the global burden of ailment and disturbs human beings in all continents across the world. Today depression is affecting more than 350 million people. Because of this, depression ends up with disorders which usually start at a very young age. It has been found that the strain of depression is more in females than males (WHO, 2008). The demand of curbing depression is on top these days. For this reason, there is an urgent need to estimate depression. The prevention of depression is an area that deserves attention. For this case, many techniques and algorithms have been developed to calculate the depression. In chapter 2, we have discussed many such existing techniques. Depression estimation is a great challenge in today's world. Therefore, in this report, we have proposed a new technique to estimate the depression using audio signals. Researchers have defined different methods to estimate the depression based on different features:

- Depression can be estimated by extracting head pose features, the Space Temporal Interesting Point (STIP) features from video. These extracted features denote the body movement and spatio - temporal changes within the image sequence.
- It can also be estimated by using facial expressions. Facial expressions can be evaluated by using video using two systems: "Manual System" and "Automatic System".
- It can also be estimated by capturing images of head pose and eye movement from

video. After that features can be extracted from the images for depression evaluation.

- It can also be estimated by asking a set of questionnaire to the people and focused on questions based on positive and negative emotions.
- It can also be estimated by using "Beavers System Model". In this model, medical history of the person is submitted through website. After that it will evaluate the chances of acquiring depression from the data.

1.4.1 Estimation of depression level from audio signals

Researchers have defined various techniques to estimate depression in audios:

- Depression can be estimated in audio signals using vocal parameters to classify between the depressed speech and suicidal speech. They used the following parameters: "Amplitude Modulation", "Formants" and "Power Distribution". These two parameters gave out the best results: "Formants" and "Power Distribution" parameters.
- Depression can also be estimated by extracting speaking rate from audios.
- Depression can also be estimated in audios by using speech processing and machine learning algorithm. Following parameters of audio can be used for depression estimation: loudness related features, sound, intensity, frequency and jitter.
- Harmonic model can also be used for depression estimation in audios. This model uses spoken words or statements from audios for depression estimation.
- Depression can also be estimated in audios by using "Canonical Correlation Analysis" based characteristic selection technique. This CCA model extracts features from audio.
- Depression can also be estimated in audios by using these two models: "Gaussian Mixture Model" and "Universal Background Model". To increase the performance of these models, they are combined at the attribute level to increase the performance of audio attributes.

Female	Male
They blame themselves.	They blame others.
They feel regretful, depressed and downhearted.	They feel irritated, angry and ego-inflated.
Feel disturbed, worried and nervous.	Feel suspicious and guarded.
They avoid disputes.	They create disputes.
Use food and friends to self-medicate.	Use alcohol, TV and sports to self-medicate.
Find it easy to talk about self-doubt and despair.	Find it weak to admit self-doubt or despair.
Feel tensed and nervous.	Feel impatient and jittery.
Feel uneasy and restless.	Troubled and uncomfortable.

Table 1.1: Difference between Male and Female Depression

This paper estimates the depression level of each speech signal using I-Vector technique and is organized in the following manner. Chapter 2 discusses some relevant related work reported in literature. These particularly focus on audio signals based depression level estimation system. Chapter 3 describes our problem statement and objective. Chapter 4 describes our proposed approach. The proposed approach analyzes audio data using I-Vector method. We have used I-Vector technique to make the method durable against various other sources in audios. The proposed approach is divided into four blocks: Silence Removal, Feature Extraction Using I-Vector, Evaluate Overlapped Audio Beats, and Evaluate Depression Using Relationship Matrix. Chapter 5 discusses experimental results. A visual analysis of the proposed algorithm is provided in this chapter. It also describes our experimental setup. This chapter also provides a brief of all the parameters which we have used in our experiment. We have taken a dataset of 20 voices and 10 voices to train our system and rest of them are used for testing purposes. The MATLAB 2013A is used along with the signal processing toolbox to simulate the desired environment. Experimental results are further discussed in this chapter. Chapter 6 describes conclusion.

Chapter 2

Related Work

In this chapter we have discussed various tools, techniques and algorithms which estimates the depression level of any person. We have discussed various parameters and characteristics to estimate the depression level.

Daniel J.France ET. Al. (2000), proposed some properties of audio speech which had been identified as possible sign to depression. Researchers had used vocal parameters to classify between the depressed speech and suicidal speech. Some experiments were performed to examine and differentiate the audio speech of male and a female sample consists of ordinary people and the samples of people carrying depression with higher risk. Female samples were used. Samples were consisted of dysthymic patients and also samples were consisted of control subjects and of major depressed patients. Male samples were also used. Samples were consisted of major depressed patients and were consisted of control subjects, and of higher risk suicidal patients. For the speech analysis they had used the following parameters: frequency (F), amplitude modulation (AM), formants and power distribution were performed on audio samples collected from members. The parameters formant and power spectral density were found to be best discriminators for male and female studies. Amplitude modulation was found to be strong discriminator for male class. Frequency parameter was found ineffective discriminator in both the cases [5].

Chokri Ben Amar ET. Al. (2012), proposed that depression is enlarging in our community. It causes sadness, affliction and trauma to patients suffering from depression. It also affects the loved ones who care about them. For preventing depression they had proposed multimodal emotion recognition system. This modal consists of detecting constant negative emotions of any individual at early stages of dysthymia. Their method was built on numerical description of psychological condition by using complex vectors. The technique gives strong statistical mechanisms for the evaluation and execution of feelings and also allows the unification of harmonious data like: sound, facial expressions, physical waves [1].

Hichem Sahli ET. Al. (2015), proposed a multimodal method for detection of depression using audio and visual cues. To increase the effectiveness of Beck Depression Inventory (BDI) except the Low Level Descriptors (LLD) features and Local Gabor Binary Pattern-Three Orthogonal Planes (LGBP-TOP) features provided by Audio Visual Emotion Challenge Workshop (AVEC2014), they had extracted some additional features to capture key behavioral changes related with depression. They had extracted speaking rate from audio and head pose features, the Space-Temporal Interesting Point (STIP) features from video. These extracted features denote the body movements and spatio-temporal changes within the image sequence. Global dynamic features were also considered which were obtained using motion history histogram (MHH), bag of words (BOW) features and vector of local aggregated descriptors (VLAD). For capturing complementary information they had evaluated two fusion systems - feature fusion scheme and model fusion scheme through local linear regression (LLR). Experiments were carried out on Depression Recognition Sub-Challenge (DSC) of AVEC2014 [2].

Linlin Chao (2015), proposed a multimodal method for depression scale prediction system built on audio-visual technique. They had extracted characteristics from audio and video and then extracted features were used and fused in feature level to describe audio visual behavior. After that they had used ”Long Short Memory Recurrent Neural Network” to encode the temporal information of the abnormal audio visual behavior. Emotion information was utilized by multi-task learning to increase the performance further. The proposed method was evaluated on the Audio-Visual Emotion Challenge (AVEC2014) dataset. The proposed method also helps in predicting depression scale [3].

Paula Lopez-Otero ET. Al. (2015), proposed a method for detection of depression from audio using I-Vector paradigm. This technique was used for depression level estimation in speech. The proposed method was grouped into three separate categories. First of all in the front-end they had removed silence in the audios by applying "Voice Activity Detector" and detected voice segments in separate steps. Features from voice recordings were extracted using the "Perceptual Linear Prediction" (PLP) coefficients which include important data about depressed voice. In speech-representation part they had trained a large number of dataset using "Universal Background Model" (UBM). Large utterances were fragmented into smaller pieces by applying a sliding window and duration of 40s and then they had applied I-Vector technique to ignore the impact of announcer and sound variability. In depression level estimation they had estimated the depression in audios using the following parameters: "Root Mean Square Error" and "Mean Absolute Error" [4].

Asgari ET. Al. (2014), work has been reported where speech processing and machine learning algorithms were used for diagnosis of depression from recordings of topics. To capture features of speech clues related with clinical depression, they had adopted traits defined in INTERSPEECH 2010 Paralinguistic Challenge using open Smile toolkit. Features were consisted of 1582 components and were divided into three different categories: In the first category loudness related features were extracted. In the second category it consists of audio related traits like sound, intensity, frequency and jitter. To measure the result of audio contents in the clinical depression they had extracted textual traits from manual transcripts. It has been seen that harmonic model improves the efficiency of diagnosis of depression from spoken words or statements [6].

Nicholas Cummins ET. Al. (2013), researchers had also worked on a multilayered model containing three distinct kernels formed using the Gaussian Mixture Model and Universal Background Model system to identify depression. It is combined at attribute level for increasing the performance of acoustic audio attributes and visual attributes given by bag of words model to identify a person's depression level using regression [7].

Chi-Chang ET. Al. (2009), researchers have explained that depression is a severe mental health problem. People suffering from depression have low feeling, depressed mood, interrupted sleep, poor spark and poor concentration. It is increasing day by day due to load and these issues become persistent and it can also destructs human's ability to carry out his or her daily task. By the year 2020, it is predicted that depression will reach at second place for all ages and for both male and female. They had proposed an online method for nuclear family which executes organized analytic technique with ”Beavers System Model”. By entering the past information of family groups through webpage, then it will evaluate the chances of acquiring depression from the information. Then it passes this data to the specialist for proper planning and also for proper treatment for sick person. They had used samples from Chung Shan Medical University of faculty and students [8].

Martin Wollmer ET. Al. (2013), proposed a fully automatic audio-visual identification method built on ”Long Short-Term Memory”. LSTM systems include knowledge about how feelings grow with the time and concluded that feeling are approximated by inspecting the best context. Large analysis on the ”Audio Visual Emotion Challenge” (AVEC2011) dataset show that how audio, verbal and video traits helps in identifying the various affective dimensions as described in the ”SEMAINE Database”. They had applied same audio traits which have been used in ”Audio Visual Emotion Challenge” method and video traits were computed through facial motion trait extractor. They compared their conclusion with the ”Audio Visual Emotion Challenge” participants dataset and LSTM Technique was proved better for recognition performance [9].

John W. Peifer ET. Al. (2008), proposed a scheme built on traits associated with the vocal region, prosodic and the specifications which are completely fetched using glottal signal. They used feature selection strategy to distinguish depressed speech by using the mixtures of characteristic domains: evaluating metrics only, evaluating metrics and vocal regions, evaluating metrics and glottal area. The mixtures of glottal and metric traits were proved to be better for the distinction of depressed speech as compared with the other combinations [10].

Li Sun ET. Al. (2013), proposed a method for the identification of depressed users in social network services by applying data mining to psychological area. In the first step sentiment analysis method was proposed by using vocabulary and human made rules to estimate the depression inclination of each micro-blog. In the second part depression identification method was constructed which was based on the proposed method and also used some traits of depressed people derived from the psychological research. After that 180 users and 3 types of divisions were used to evaluate the model and whose precisions were around 80 per cent. The importance of each trait was analyzed. And in the last an application was developed for monitoring mental health online [11].

Gordon Parker ET. Al. (2013), proposed a method for identification of depression using "Affective Sensing Technology" and focused on audio traits. First of all they had classified the general properties of clinical depression by using spontaneous speech as compared to the read speech, because it will give more efficient results. In the second part they had used some audio traits which were very powerful and also provides efficient categorization conclusion in "Spontaneous" and "Read" speech. In the third part they had used thin slicing method in which small parts of speech were used to perform similarly and if it was not efficient then whole part of speech was used. After that they had analyzed and compared the identification results of audio traits on actual time clinical file. Dataset was consisted of depressed samples and control subjects by using support vector machine for classification and they found that spontaneous speech had more variability as compared to the read speech which increases the identification rate of depression. Some powerful features were helpful in classifying between spontaneous and depressed speech like jitter, shimmer, tone and energy [12].

Julien Epps ET. Al. (2013), proposed a method for examining the properties of depressed speech for the motivation of automatic classification by examining the effect of various audio features on the classification results. They had examined the voice and the unvoiced part and the mixed speech part in order to gain an efficient solution for depressed speech and also to connect the gap between physiological and affective computing studies. This solution will ultimately provides an effective method that will support the doctors in identification and also helps in estimation of depression. A relation has been formed between the tests of depressed speech using "Gaussian Mixture Model" and "Support Vector Machine" and Variance analysis is utilized to categorize its features. They have used speech utterances of depressed patients for feature extraction and classification [13].

James R. Williamson ET. Al. (2013), researchers had utilized the results that reflect variation in vocal region movement related with "Major Depression Disorder". Researchers had also examined variation in association which occurs at various points covering formant distribution and also across delta mel-cepstrum signals. Both traits give measures of vocal tract articulation and also reduce the effect of slowly varying channel. With the help of these two additional features using the Audio Visual Emotion Challenge (AVEC2013) depression dataset, researchers have created a blueprint of "Gaussian Mixture Model" (GMM) based on "Multivariate Regression Scheme. It is also referred as "Gaussian Staircase Regression [14].

Jiping Li ET. Al. (2011), researchers have explained that the identification and suppression of small traits from solid B-rep models plays a vital role in creating the qualified mesh for investigation. They proposed an automatic algorithm identifying small depression traits for the justification of mesh generation. The alteration of the algorithm lies in the identification capability for filleted depression trait [15].

Heysem Kaya ET. Al. (2014), proposed a method for depression identification from speech by making use of Canonical Correlation Analysis (CCA) based feature selection technique. This technique is commonly used in multimodal feature extraction. This technique can also be used as a feature selector. They had introduced several novel methods of CCA based methods in which they shown their relations to previous work. They had tested their proposed technique on the Audio Visual Emotion Challenge (AVEC2013) dataset under the ACM MM 2013 challenge protocol to estimate the depression level [16].

Laura Docio-Fernandez ET. Al. (2014), researchers have explained that soft-biometry consists of botanical characteristics which are not enough for the identification of any individual yet it narrows the search space. Evidence of mental condition of an individual can be considered as a soft biometric feature because it gives enough data for the recognition of an individual. For the categorization of healthy and depressed speech various methods have been utilized. Yet the algorithm, traits and execution evaluation makes it very tough to conclude regarding which technique, traits will be best for the given task. They had also explained that they had provided explanation regarding audio traits which were used for the categorization of depressed voice and was studied on "Audio Visual Emotion Challenge" (AVEC2013) dataset. They proposed a method where speech files were projected, and then segmented into an entire area and depression level can be evaluated using predicted information [17].

Nicholas Cummins ET. A.l. (2013), researchers had investigated the theory that essential depression based data can be fetched inside the covariance frame of a "Gaussian Mixture Model" of documented speech. Important negative correlations discovered between a speaker's average weighted variance and their amount of depression holds this hypothesis. Another evidence is given by the comparing the classification accuracies from seven distinct GMM-UBM systems, each of which is formed by changing dissimilar attribute combinations during MAP adaption. This method shows that variance-only adaptation beats the de facto standard mean-only adaptation when categorizing both the severity and presence of depression [18].

Sharifa Alghowinem ET. Al. (2013), researchers had used clinically proved dataset of real world. "Gaussian Mixture Models" (GMM), "Support Vector Machines" (SVM) and "Multilayer Perceptron Neural Networks" (MLP) were compared using a powerful classifier from computing literature and also introduced a latest classifier known as Hierarchical Fuzzy Signature (HFS) classifier. GMM and SVM were proved to be best among all the other classifiers and provides best results as compared with other classifiers. Researchers compared characteristics, marks and decision fusion executed well on these three classifiers: Gaussian Mixture Models, Hierarchical Fuzzy Signature and Multilayer Perceptron Neural Networks, while decision fusion performed well for Support Vector Machine. Only feature fusion executed badly than other fusions techniques in this research. They also found that pitch, RMS, and intensity were the audio traits which executed well for identifying depression in the corresponding dataset [19].

Roland Goecke ET. Al. (2013), researchers had developed a method named "Affective Sensing Technique" which holds doctors in their identification and observation of medical depression. They had also evaluated the performance of eye movement characteristics and those eye movement characteristics were fetched from the face videos by utilizing "Active Appearance Models" for a dyadic categorization work. They also found that low level traits of eye movement gave high accuracy by utilizing a hybrid classifier of GMM and "Support Vector Machines" (SVM), and statistical classifier gave best results when SVM classifier was used over the complete interview. The difference between positive and negative emotions were also investigated and also evaluated the performance in gender-dependent and gender-independent modes but the wink speed was not extraordinary among the healthy people and depressed people. They also explored that the range between the eyelids was very less and the time of wink was very large in depressed people, which depicts tiredness or eye contact avoidance [20].

Michael Wagner ET. Al. (2012), proposed a method for diagnostic aid supporting the doctors by using the affective sensing technique. They focused on acoustic and statistical traits from spontaneous speech. Their research evaluated the change in expressing negative and positive feelings in depressed people and healthy people and also investigated that the initial gender categorization increases the identification rate. Spontaneous speech data was consisted of 30 subjects of depressed people and healthy people was evaluated, and focused on questions based on positive and negative emotions. Hidden Markov Model and Gaussian Mixture were used for the categorization. They discovered that MFCC, strength and power features produced maximum identification rates when both female and male subjects were evaluated simultaneously. First time when the dataset was divided by the gender, shimmer and Root Mean Square Error (RMSE) gave the highest identification rate in females and audio quality for men. Using temporal characteristics, response time along with average syllable duration were very large in depressed people. The articulation rate was very high in control subjects [21].

Michael Break Spear ET. Al. (2013), work has been reported where an affective sensing technique was used that holds doctor in their identification and observation of medical depression. They had used a 3D face model which was used to evaluate the ability of head pose and motion traits. In a dyadic categorization task, they had modeled low-level and statistical functional traits for an SVM classifier utilizing real-world clinically proved information. The head position and motion can be used as a supportive function at identifying depression and corresponding identification rate was very good and gave best results which defines that the head position and motion hold successful clues in identifying depressed people [22].

Jyoti Joshi ET. Al. (2013), researchers had developed an identification method for depression using affective sensing technology. They proposed a multimodal method comprised of audio and video together for depression recognition. They had utilized the theory in which the auditory and visual human interaction accompanies each other, which was very famously known in auditory-visible audio processing; they had examined the theory for depression evaluation. For the evaluation of visual data, intra-facial muscle motions and the motions of the shoulders and head were evaluated with the help of calculating spatiotemporal interest points. A collection of video characteristics and a collection of audio features were produced independently. They had compared fusion methods at characteristics level and score level and decision level. Tests were executed on an age and gender. The medical dataset was consisted of depressed victims and good condition people [23].

Stefan Scherer ET. Al. (2014), researchers had evaluated that depression alters vocal timing of both patients and medical interrogators but were mixed in respect of audio characteristics. They had additionally analyzed middle aged adults with "Major Depression Disorder" and their corresponding medical interrogators were analyzed. Candidates were questioned for depression asperity on up to four instants over a several week period utilizing the Hamilton Rating Scale for Depression (HRSD), which used to be a criterion aspect for depression severity in medical trials. Audio traits were fetched for both candidates and interrogators by utilizing COVAREP Toolbox. Absent data was occurred due to lost appointments, technical complications, or not enough vocal samples. Data from patients and their interrogators met benchmark and were presented for evaluation to correlate between higher depression and lower depression asperity. Audio traits for candidates changed between male and female as expected, and failed to vary with depression severity for candidates. For interrogators, audio traits were strongly differs with seriousness of interviewer's depression [24].

Rahul Gupta ET. Al. (2014), proposed a multimodal signal processing techniques to confront two problems using information from human-computer communications. They had developed different frameworks for vaticinator of depression states and affective dimensions, implementing by using various techniques for joining the multimodal data. The presented depression prediction method used a characteristics choosing method dependent on audio, video, and linguistic clues to foresee depression scores for every session. They had used multiple techniques trained on audio and visual clues to foresee the affective dimensions in continuous time. Their affect identification method accounts for circumstances at the time the frame wise inference and implements a linear fusion of results from the audio and visual systems. For both difficulties, their proposed systems produced much better results than the video-feature based baseline techniques. They had analyzed the role played by every modality in calculating the target variable and produce analytical insights [25].

Kuan Ee Brian Ooi ET. Al. (2012), researchers have evaluated the effectiveness of Teager Energy and glottal characteristics in an ancient prediction of "Major Depression Disorder" in youngsters utilizing a completely fully automated speech inspection and categorization method. The estimation was attained by a dyadic categorization of recordings of the audio taken from youngsters who produced "Major Depression Disease" within two years after these recordings were formed and the samples of youngsters who did not produced "Major Depression Disease" within the same time period [26].

Lu-Shih Alex Lo ET. Al. (2011), researchers had analyzed a technique that the feasibility of using face pictures to decide if a good condition person is likely to produce medical depression within the coming 1-2 years. The vaticination approach utilizes a special categorization method of training and testing. The class models were identified with the help of picture information from youngsters who were either at certainty or uncertainty of depression. The risk element was validated from two years of follow-up dataset. Two characteristics extraction methods were distinguished, the "Eigenface" (PCA) features and the "Fisherface" (PCA+LDA)traits. The Nearest Neighbor (NN) categorization was executed using person independent and person dependent methods [27].

Stefan Scherer ET. Al. (2013), researchers had examined that in the automatic evaluation of emotional situations like depression, "Post-Traumatic Stress Disorder", gender plays a vital role. They had also determined a directly portrayed and perceptive set of predictive indicators and choose from three common groups of nonverbal behaviors which are "Affect", "Expression Variability" and "Motor Variability". They included a semi-structured virtual human questioning dataset which included video recorded communications of people [28].

Jeffrey M. Girard ET. Al. (2013), researchers have searched the connection between alteration over time in seriousness of depression characteristics and face expression. The depressed candidates were accompanied over the course of medical treatment and videos were recorded at the time of a sequence of medical evaluation. Face expressions were examined by the video using two systems: physical system and automatic system. For both the systems coding was totally comprised of FACS action units, and it displayed the same consequences for alteration over a span in depression extremity. For both the systems, when expression extremity was very higher, candidates form extra face expressions related with disrespect, become very unhappy, and the smiles that happened were more likely to be associated with face actions related with contempt. The produced results were constant with the "social risk hypothesis" of depression. According to social risk hypothesis, when expressions were very critical, depressed candidates extracted from other patients so as to save themselves from predicted rejection, scorn, and social exclusion. Participants produce large signals indicating an eagerness to associate, as their symptoms grow dim. The judgment that automated facial expression analysis was both constant with man-made coding and reproduced the similar pattern of depression consequences suggests that automatic facial expression examination may be available for use in behavioral and medical science [29].

Louis-Philippe Morency ET. Al. (2013), researchers have investigated the ability of automatic nonverbal behavior descriptors to recognize the signs of psychological traits like depression, nervousness, and "Post-Traumatic Stress Disorder". They improve and verify the recent techniques, mainly dependent on qualitative manual observation, with automatic behavior descriptors. They proposed four nonverbal behavior descriptors which can be automatically calculated from visual signals. They developed a latest file called the "Distress Assessment Interview Corpus" in which binary interactions were contained between a confederate interrogator and a paid member. The estimation on this data files presents association of our automatic behavior descriptors with particular psychological traits along with a generic distress measure. Their investigation also contains a deeper research of self adaptor and fidgeting behaviors dependent on detailed observation of where these behaviors take place [30].

Chapter 3

Problem Statement and Objectives

3.1 Problem Statement

This dissertation has focused on speaker dependence of an I-Vector based depression level estimation system. I-Vector based depression level estimation technique has better performance than majority of existing techniques. But the effect of the silence has not been removed in the input signals while depression level estimation. Therefore this dissertation will propose a silence removal and I-Vector based depression level estimation technique.

3.2 Objectives

- To design and implement and propose silence removal from audio signals.
- To propose an I-Vector based depression level estimation method.
- To evaluate the performance of the I-Vector based depression level estimation technique.

Chapter 4

Estimation of Depression Level in Audios Using I-Vector Technique

The technique used in our analysis depicts audio data using I-Vector method. We have used I-Vector technique to make the method durable against various other sources in the audios. First of all the audios will be loaded into the workspace. Our proposed method is divided into four different parts: In the first part, we have trained the audio signals and then silence has been removed from the audio signals. In the second part, features were extracted from audios using I-Vector. In the third part, split overlapping function is applied to evaluate the overlapped audio beats. In the fourth part, we have evaluated depression using relationship matrix. Figure 4.1 is showing the block diagram of depression estimation of audio signals using I-Vector method. Table 4 refers to nomenclature of metrics used in equations.

4.1 Silence Removal

In this part we have removed silence from the audio signals using the following equations as given below. In equation 4.1, we calculated the sum of amplitude frequency of signal. In equation 4.2 and 4.3, we calculated the minimum amplitude frequency of signal to remove silence from the signal.

$$amp_f = \sum(abs(spt_{ovp}(ado([1 - 0.9375], 1, x), adoLen, adoInc)), 2) \tag{4.1}$$

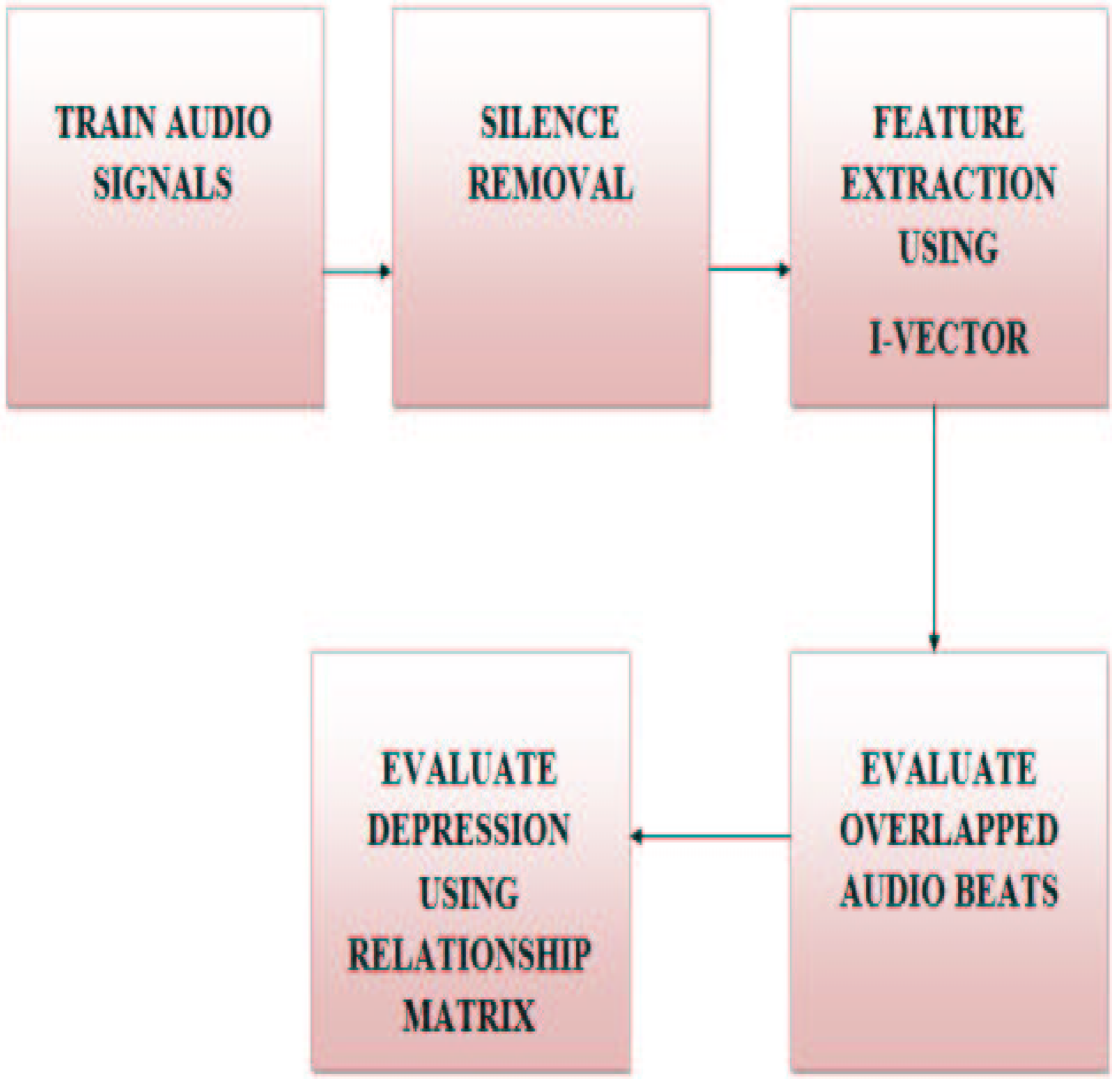

Figure 4.1: Block Diagram of Depression Estimation

Metric	Full Name
Amp	Amplitude
F	Frequency
Spt	Split
Ovp	Overlapping
Ado	Audio
Len	Length
Inc	Increment
Abs	Absolute
Min	Minimum
Max	Maximum
Coef	Coefficient
Fft	Fast Fourier transform
Dist	Distance
Corr	Correlation
Tp	True Positive
Tn	True Negative
Fp	False Positive
Fn	False Negative

Table 4.1: Nomenclature of metrics used in equations

$$amp_{f1} = min(amp_{f1}, \frac{max(amp_f)}{4}) \tag{4.2}$$

$$amp_{f2} = min(amp_{f2}, \frac{max(amp_f)}{8}) \tag{4.3}$$

Algorithm for silence removal
$if amp_{fq(n)} > amp_{fq1}$
x1 = max(n-count-1,1);
$elseif amp_{fq}(n) > amp_{fq2}$
$OR zcount(n) > zcount2$
silent = silent+1;
$if silent < peak_{silent}$
silent = 0;
end

Table 4.2: Algorithm for removing silence from audio signals

4.2 Feature Extraction Using I-Vector

In the second part we have extracted features from audio signals using I-Vector method using the following algorithm:

a. Initialize vector variable with Melvectorm function using equation 4.4:

$$vector = melvectorm(audio) \tag{4.4}$$

b. Calculate frequency to bit-ratio and store it in random variable $l_r$using equation 4.5.

$$l_r = \frac{\lg \frac{f_{zero}+fh}{f_{zero}+f1}}{p+1} \tag{4.5}$$

c. Convert l_r value to fast fourier transform bin numbers using equation 4.6, 4.7, 4.8, 4.9 and 4.10.

$$b1 = n \times ((f_{zero} + f1) \times exp([01pp + 1] \times l_r) - f_{zero}) \tag{4.6}$$

$$p_f = \frac{\lg(f_{zero} + \frac{b2:b3}{n} f_{zero} + f1}{l_r} \tag{4.7}$$

$$r = [ones(1, b2)fpfp + 1P \times ones(1, frq_{n2} - b3)] \tag{4.8}$$

$$c = [1 : b3 + 1b2 + 1 : frq_{n2} + 1] \tag{4.9}$$

$$v = 2 \times [0.5ones(1, b2 - 1)1 - p_f + f_p p_f - f_p ones(1, frq_{n2} - b3 - 1)0.5] \tag{4.10}$$

d. Using equation 4.11 and 4.12 we calculated the value of melvectorm function.

$$vector = 1 - \frac{0.92}{1.08} \times cos(v \times \frac{pi}{2}) \tag{4.11}$$

$$vector = \frac{vector}{max(vector(:))} \tag{4.12}$$

e. In equation 4.13 and 4.14 we calculated the vector value of signal and store it in variable w.

$$w = 1 + 6 \times sin(pi \times \frac{[1:12]}{12}) \tag{4.13}$$

$$w = \frac{w}{max(w)} \tag{4.14}$$

f. Store the result.

4.3 Evaluate Overlapped Audio Beats

In this part we have evaluated overlapped audio beats using split overlapping function using the following algorithm:

- Read the audio data.
- Multiply the audio data with the hamming distance.
- Apply the fast fourier transform function on the above data.
- Calculate the distance vector and store it in distance vector variable.
- Calculate correlation matrix.
- Evaluate the overlapped audio beats using split overlapping function.
- Store the result.

4.4 Evaluate Depression Using Relationship Matrix

In this part we have evaluated depression using the equation 4.15:

$$relation(i,j) = \sum(\sqrt{(t(i,:) - r(j,:))}) \tag{4.15}$$

Here t is the trained signal and r is the input signal. For given signal t, if the matched value of r signal is high then depression will be the depression value stored in the trained data set.

Chapter 5

Experimental Results

In this chapter firstly we have discussed about visual analysis in audio signals. We have matched the actual signal with the test signal. After that we have described about experimental setup which we have used in our analysis. This chapter contains all the information about the parameters which we have used in our experiment. After that we have discussed experimental results. This chapter contains the overall result of the experiment.

5.1 Visual Analysis

Figure 5.1, 5.2 and 5.3 are showing the input or actual signal as well as the test or matched signal and both the signals are very similar. After that we will calculate the value of accuracy, specificity, peak signal to noise ratio, f-measure and balanced classification rate for test signal. After calculating the values of these parameters we will estimate the depression level in each signal.

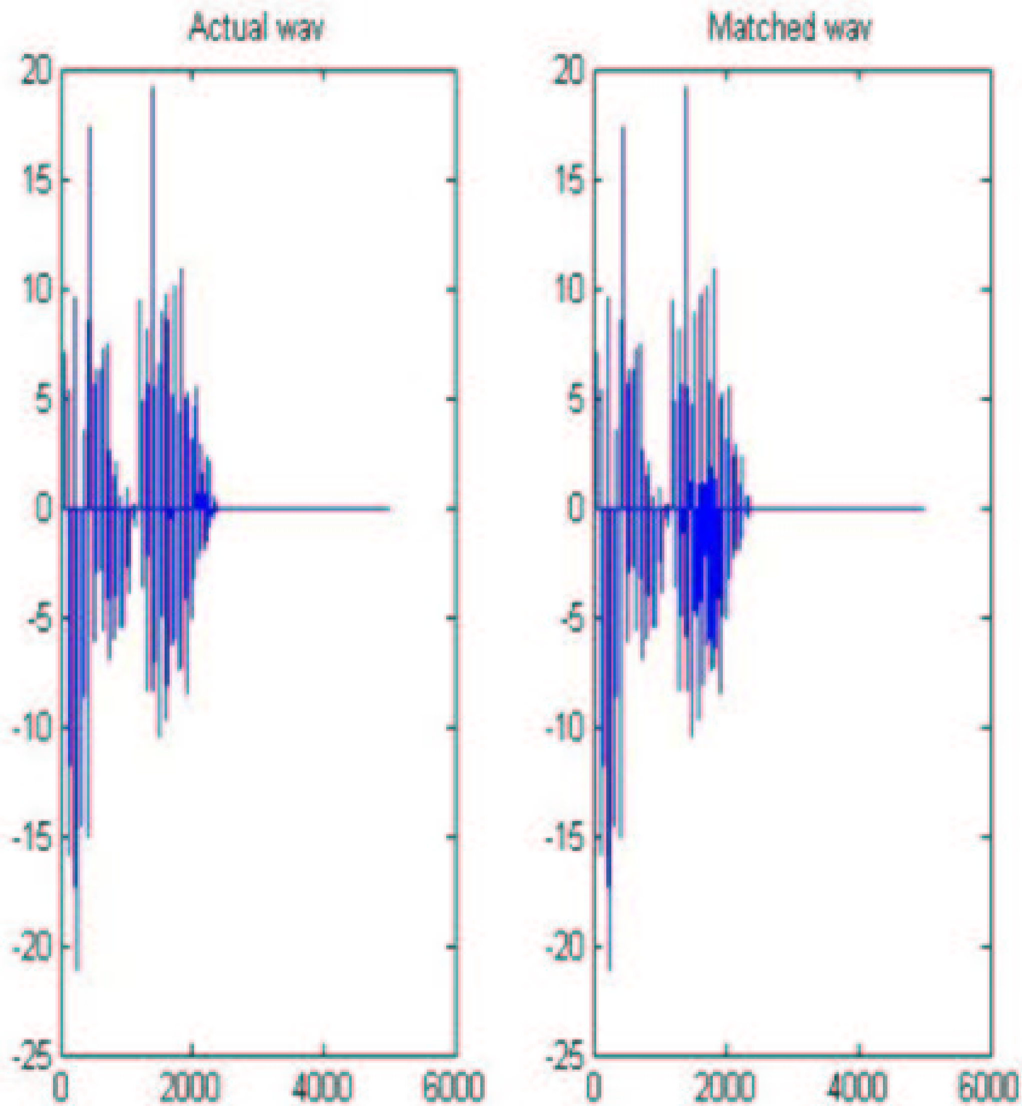

Figure 5.1: Diagram of actual waveform and matched waveform

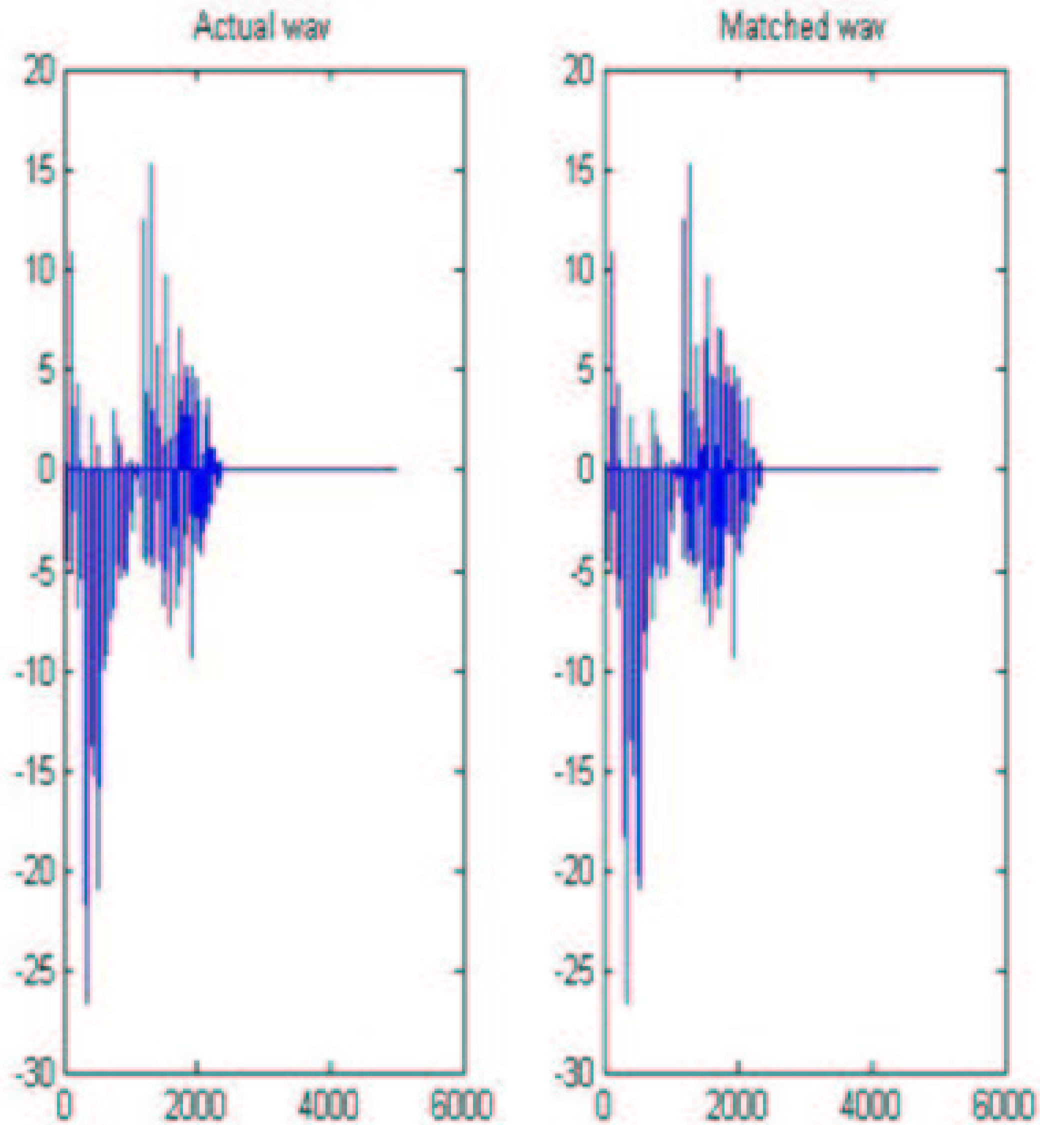

Figure 5.2: Diagram of actual waveform and matched waveform

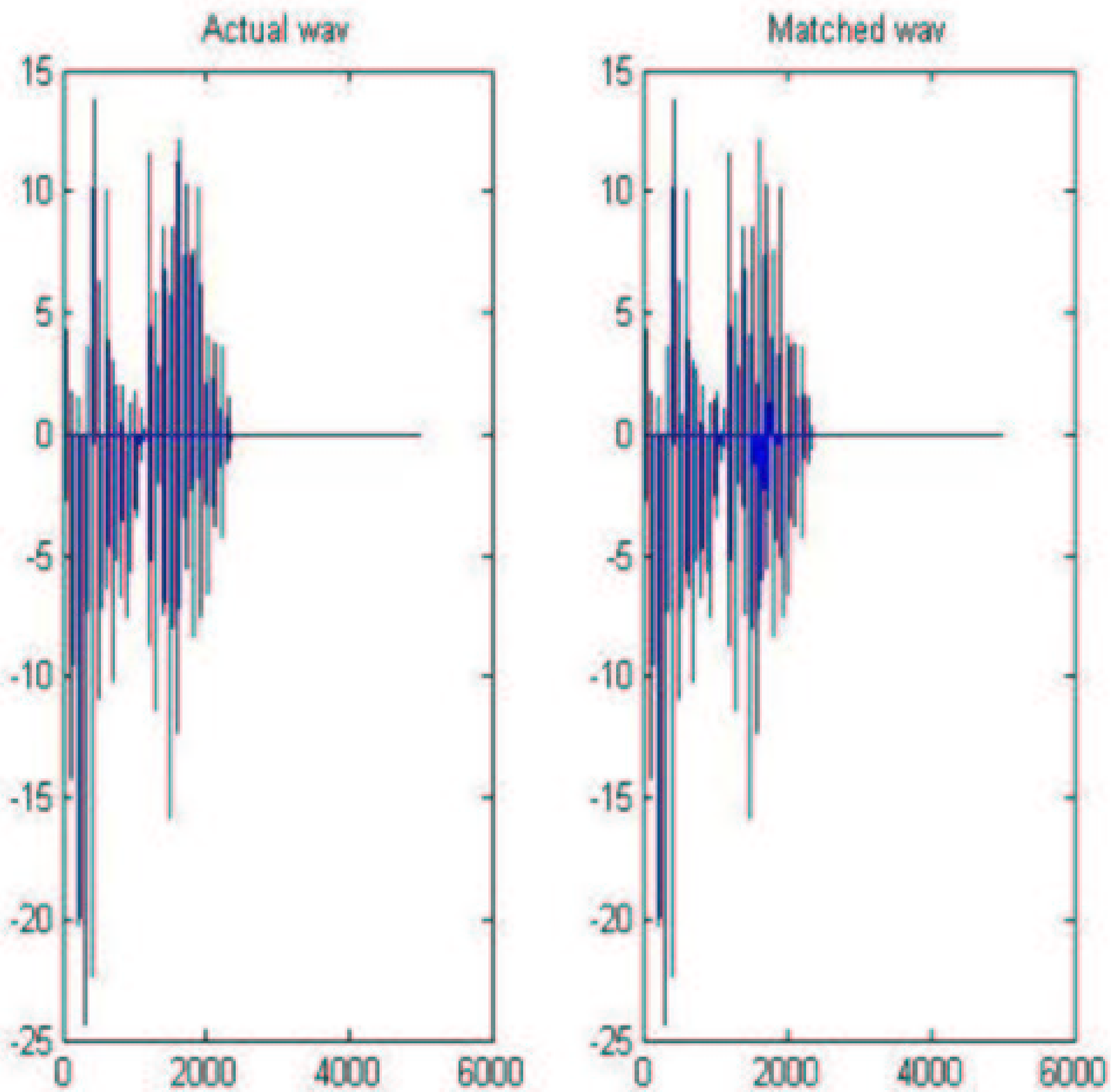

Figure 5.3: Diagram of actual waveform and matched waveform

5.2 Experimental Setup

The experimental setup contains the overall experimental detail of the I-Vector based depression level estimation technique. The MATLAB 2013A is used along with the signal processing toolbox to simulate the desired environment. We have taken a dataset of 20 voices and used 10 voices to train our system and rest of them are used for training purposes. Table 5.1 is showing the required detail of parameters which are required to simulate the I-Vector based depression estimation. However, I-Vector based depression level estimation is not limited to the given set of audio signals.

Parameters	Values
Number of trained signals	10
Number of test signals	10
Audio length	240
Audio Inc	80
Amp_fq1	10
Amp_fq2	2
Zcount1	10
Zcount2	5
$Peak_silent$	8
Min_size	15
Processed	0
Count	0
Silent	0

Table 5.1: Parameters for I-Vector based depression level estimation technique

5.3 Experimental Results

This section contains the performance evaluation of the I-Vector based depression level estimation system. Five different well-known quality metrics are considered. Various metrics are defined as follow. Accuracy is defined as the mean of sensitivity and specificity and we have calculated the accuracy for the signals by using equation 5.1:

$$Accuracy = \sum \frac{tp + tn}{tp + tn + fp + fn} \tag{5.1}$$

Here, TP is defined as the number of true positives, TN is defined as the number of true negatives, FP is defined as the number of false positives, FN is defined as the number of false negatives. Specificity is defined as the true negative rate and we have calculated the specificity of the signals by using equation 5.2:

$$Specificity = \sum \frac{tn}{tn + fp} \tag{5.2}$$

Here, TN is number of true negatives and FP is false positives. Peak signal to noise ratio calculates the ratio between the maximum attainable value of a signal and the value of degrading sound that changes the accuracy of its representation. We have calculated the PSNR by using equation 5.3:

$$PSNR = 10 \times \lg 10 \frac{1}{error} \tag{5.3}$$

Here, error calculates the error rate and error rate is calculated by using equation 5.4:

$$ErrorRate = \sum \frac{fp + fn}{tp + tn + fp + fn} \tag{5.4}$$

We have calculated the balanced classification rate for the signals by using equation 5.5:

$$BCR = 0.5 \times (sensitivity + specificity) \tag{5.5}$$

Here, sensitivity is calculated by using equation 5.6:

$$Sensitivity = \sum \frac{tp}{tp + fn} \tag{5.6}$$

Signal	Accuracy	Specificity	PSNR	BCR	F-Measure	Depression Level
1	0.9584	0.6837	13.8510	0.8296	97.7868	20
2	0.9548	0.5833	13.4872	0.7784	97.6216	30
3	0.9580	0.6433	13.8091	0.8112	97.7645	40
4	0.9482	0.4508	12.8904	0.7134	97.2768	50
5	0.9572	0.6216	13.7263	0.8000	97.7297	60
6	0.9456	0.5888	12.6761	0.7794	97.0934	70
7	0.9514	0.5760	13.1695	0.7750	97.4256	80
8	0.9309	0.4689	11.6241	0.7158	96.3034	10
9	0.9632	0.6846	14.3890	0.8327	98.0459	20
10	0.9321	0.5328	11.7005	0.7496	96.3227	30

Table 5.2: Estimation of Depression Level

F-Measure is defined as the harmonic mean between precision and recall and we have calculated the F-Measure for the signals by using equation 5.7:

$$F - Measure = 100 \times 2 \times \frac{precision \times recall}{precision + recall} \tag{5.7}$$

Here, precision is calculated by using equation 5.8:

$$Precision = \sum \frac{tp}{tp + fp} \tag{5.8}$$

Here, recall is calculated by using equation 5.9:

$$Recall = \sum \frac{tp}{fn + tp} \tag{5.9}$$

Here the values of all these parameters have been calculated in MATLAB 2013A. And after calculating the values of these parameters we have calculated the depression level in each speech signal. As in prior Accuracy, Specificity, Peak Signal to Noise Ratio(PSNR), Balanced Classification Rate (BCR) and F-Measure need to be maximized and the mean value for accuracy is 0.949, for specificity is 0.583, for PSNR is 13.132, for BCR is 0.778, for F-Measure is 97.33. Therefore the overall result clearly indicates the good performance of the I-Vector based depression level estimation system as shown in the given Table 5.2.

Chapter 6

Conclusion

In this dissertation we have focused on depression level estimation system using I-Vector based technique. The use of I-Vector technique has ability to reduce the effect of silence from audio signals which interrupts the accuracy in earlier depression level estimation methods. Silence has been removed from the audio signals. Then features have been extracted from audio using I-Vector. Then split overlapping function has been applied to evaluate the overlapped audio beats. Then depression is evaluated in audio signals using relationship matrix. Then the depression level of each speaker is estimated. Thus it has more accuracy than existing techniques. We have got 93 per cent accuracy, which describes effectiveness of I-Vector based depression level estimation system. Thus, it is more practical to use in real time applications. In future we will use fuzzy membership functions to estimate the depression level and we will compare it with I-Vector technique to get better results.

Publication

1. Barkha Rani , "I-Vector Based Depression Level Estimation Technique", in Recent Trends in Electronics, Information and Communication Technology (RTEICT), 2016 IEEE International Conference on, Bangalore, India, IEEE, 2016. (Status: Accepted)

Video

1. $https://youtu.be/oE4Wvo_aLNM$

Bibliography

[1] Meftah, Imen Tayari, Nhan Le Thanh, and Chokri Ben Amar. "Detecting depression using multimodal approach of emotion recognition." In Complex Systems (ICCS), 2012 International Conference on, pp. 1-6. IEEE, 2012.

[2] He, Lang, Dongmei Jiang, and Hichem Sahli. "Multimodal depression recognition with dynamic visual and audio cues." In Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on, pp. 260-266. IEEE, 2015.

[3] Chao, Linlin, Jianhua Tao, Minghao Yang, and Ya Li. "Multi task sequence learning for depression scale prediction from video." In Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on, pp. 526-531. IEEE, 2015.

[4] Lopez-Otero, Paula, Laura Docio-Fernandez, and Carmen Garcia-Mateo. "Assessing speaker independence on a speech-based depression level estimation system." Pattern Recognition Letters 68 (2015): 343-350.

[5] France, Daniel J., Richard G. Shiavi, Stephen Silverman, Marilyn Silverman, and D. Mitchell Wilkes. "Acoustical properties of speech as indicators of depression and suicidal risk." Biomedical Engineering, IEEE Transactions on 47, no. 7 (2000): 829-837.

[6] Asgari, Meysam, Izhak Shafran, and Lisa B. Sheeber. "Inferring clinical depression from speech and spoken utterances." In Machine Learning for Signal Processing (MLSP), 2014 IEEE International Workshop on, pp. 1-5. IEEE, 2014.

[7] Cummins, Nicholas, Jyoti Joshi, Abhinav Dhall, Vidhyasaharan Sethu, Roland Goecke, and Julien Epps. "Diagnosis of depression by behavioural signals: a multimodal approach."

In Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge, pp. 11-20. ACM, 2013.

[8] Chang, Chi-Chang, Ruey-Shin Chen, Kuo-Hsiung Liao, and Pei-Ran Sun. "Estimation of Depression Risky of Family Health Tree System-Using the Family Functioning Scale." In Management and Service Science, 2009. MASS'09. International Conference on, pp. 1-4. IEEE, 2009.

[9] Wllmer, Martin, Moritz Kaiser, Florian Eyben, Bjrn Schuller, and Gerhard Rigoll. "LSTM-Modeling of continuous emotions in an audiovisual affect recognition framework." Image and Vision Computing 31, no. 2 (2013): 153-163.

[10] Moore, Elliot, Mark A. Clements, John W. Peifer, and Lydia Weisser. "Critical analysis of the impact of glottal features in the classification of clinical depression in speech." Biomedical Engineering, IEEE Transactions on 55, no. 1 (2008): 96-107.

[11] Wang, Xinyu, Chunhong Zhang, Yang Ji, Li Sun, Leijia Wu, and Zhana Bao. "A depression detection model based on sentiment analysis in micro-blog social network." In Trends and Applications in Knowledge Discovery and Data Mining, pp. 201-213. Springer Berlin Heidelberg, 2013.

[12] Alghowinem, Sharifa, Roland Goecke, Michael Wagner, Julien Epps, Michael Breakspear, and Gordon Parker. "Detecting depression: a comparison between spontaneous and read speech." In Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, pp. 7547-7551. IEEE, 2013.

[13] Alghowinem, Sharifa, Roland Goecke, Michael Wagner, Julien Epps, Gordon Parker, and Michael Breakspear. "Characterising depressed speech for classification." In Interspeech, pp. 2534-2538. 2013.

[14] Williamson, James R., Thomas F. Quatieri, Brian S. Helfer, Rachelle Horwitz, Bea Yu, and Daryush D. Mehta. "Vocal biomarkers of depression based on motor incoordination." In Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge, pp. 41-48. ACM, 2013.

[15] Li, Jiping, Lu Sun, Jianjun Peng, Jiang Du, and Lin Fan. "Automatic small depression feature recognition from solid B-rep models for meshing." InElectrical and Control Engineering (ICECE), 2011 International Conference on, pp. 4386-4389. IEEE, 2011.

[16] Kaya, Heysem, Florian Eyben, Albert Ali Salah, and Bjorn Schuller. "CCA based feature selection with application to continuous depression recognition from acoustic speech features." In Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on, pp. 3729-3733. IEEE, 2014.

[17] Lopez-Otero, Paula, Laura Docio-Fernandez, and Carmen Garcia-Mateo. "A study of acoustic features for the classification of depressed speech." InInformation and Communication Technology, Electronics and Microelectronics (MIPRO), 2014 37th International Convention on, pp. 1331-1335. IEEE, 2014.

[18] Cummins, Nicholas, Julien Epps, Vidhyasaharan Sethu, Michael Breakspear, and Roland Goecke. "Modeling spectral variability for the classification of depressed speech." In Interspeech, pp. 857-861. 2013.

[19] Alghowinem, Sharifa, Roland Goecke, Michael Wagner, Julien Epps, Tom Gedeon, Michael Breakspear, and Gordon Parker. "A comparative study of different classifiers for detecting depression from spontaneous speech." InAcoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, pp. 8022-8026. IEEE, 2013.

[20] Alghowinem, Sharifa, Roland Goecke, Michael Wagner, Gordon Parker, and Michael Breakspear. "Eye movement analysis for depression detection." InImage Processing (ICIP), 2013 20th IEEE International Conference on, pp. 4220-4224. IEEE, 2013.

[21] Alghowinem, Sharifa, Roland Goecke, Michael Wagner, Julien Epps, Michael Breakspear, and Gordon Parker. "From Joyous to Clinically Depressed: Mood Detection Using Spontaneous Speech." In FLAIRS Conference. 2012.

[22] Alghowinem, Sharifa, Roland Goecke, Michael Wagner, Gordon Parkerx, and Michael Breakspear. "Head pose and movement analysis as an indicator of depression." In Affective

Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on, pp. 283-288. IEEE, 2013.

[23] Joshi, Jyoti, Roland Goecke, Sharifa Alghowinem, Abhinav Dhall, Michael Wagner, Julien Epps, Gordon Parker, and Michael Breakspear. "Multimodal assistive technologies for depression diagnosis and monitoring." Journal on MultiModal User Interfaces 7, no. 3 (2013): 217-228.

[24] Scherer, Stefan, Zakia Hammal, Ying Yang, Louis-Philippe Morency, and Jeffrey F. Cohn. "Dyadic behavior analysis in depression severity assessment interviews." In Proceedings of the 16th International Conference on Multimodal Interaction, pp. 112-119. ACM, 2014.

[25] Gupta, Rahul, Nikolaos Malandrakis, Bo Xiao, Tanaya Guha, Maarten Van Segbroeck, Matthew Black, Alexandros Potamianos, and Shrikanth Narayanan. "Multimodal prediction of affective dimensions and depression in human-computer interactions." In Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge, pp. 33-40. ACM, 2014.

[26] Ooi, Kuan Ee Brian, Lu-Shih Alex Low, Margaret Lech, and Nicholas Allen. "Early prediction of major depression in adolescents using glottal wave characteristics and teager energy parameters." In Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, pp. 4613-4616. IEEE, 2012.

[27] Ooi, Kuan Ee Brian, Lu-Shih Alex Low, Margaret Lech, and Nicholas Allen. "Prediction of clinical depression in adolescents using facial image analaysis." In WIAMIS 2011: 12th International Workshop on Image Analysis for Multimedia Interactive Services, Delft, The Netherlands, April 13-15, 2011. TU Delft; EWI; MM; PRB, 2011.

[28] Stratou, Giota, Stefan Scherer, Jonathan Gratch, and Louis-Philippe Morency. "Automatic nonverbal behavior indicators of depression and PTSD: Exploring gender differences." In Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on, pp. 147-152. IEEE, 2013.

[29] Girard, Jeffrey M., Jeffrey F. Cohn, Mohammad H. Mahoor, Seyedmohammad Mavadati, and Dean P. Rosenwald. "Social risk and depression: Evidence from manual and automatic facial expression analysis." In Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on, pp. 1-8. IEEE, 2013.

[30] Scherer, Stefan, Giota Stratou, Mohamed Mahmoud, Jill Boberg, Jonathan Gratch, Alessandro Rizzo, and Louis-Philippe Morency. "Automatic behavior descriptors for psychological disorder analysis." In Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on, spp. 1-8. IEEE, 2013.

Barkha_thesis_801433006

ORIGINALITY REPORT

7%	3%	4%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Scherer, Stefan, Giota Stratou, Marwa Mahmoud, Jill Boberg, Jonathan Gratch, Albert Rizzo, and Louis-Philippe Morency. "Automatic behavior descriptors for psychological disorder analysis", 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), 2013. Publication	1%
2	www.ee.imperial.ac.uk Internet Source	1%
3	France, Daniel J. Shiavi, Richard G. Sil. "Acoustical Properties of Speech as Indicators of Depression and Suicidal Risk.", IEEE Transactions on Biomedical Engineer, July 2000 Issue Publication	<1%
4	Submitted to University of Bristol Student Paper	<1%
5	www.binfieldsurgery.co.uk Internet Source	<1%

6	Lecture Notes in Computer Science, 2013. Publication	<1%
7	Submitted to 8779 Student Paper	<1%
8	www.depressionhurts.ca Internet Source	<1%
9	www.helpguide.org Internet Source	<1%
10	Wang, Xinyu, Chunhong Zhang, Yang Ji, Li Sun, Leijia Wu, and Zhana Bao. "A Depression Detection Model Based on Sentiment Analysis in Micro-blog Social Network", Lecture Notes in Computer Science, 2013. Publication	<1%
11	www.anxiety-central.com Internet Source	<1%
12	Stratou, Giota, Stefan Scherer, Jonathan Gratch, and Louis-Philippe Morency. "Automatic Nonverbal Behavior Indicators of Depression and PTSD: Exploring Gender Differences", 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, 2013. Publication	<1%
13	Gomathy, B., S. M. Ramesh, and A.	<1%

Shanmugam. "A Hybrid Model for Enhanced Prediction of Medical Diagnosis Based on Discriminative Rule Framing and Correlated Framework Approaches", International Journal of Uncertainty Fuzziness and Knowledge-Based Systems, 2015.
Publication

---

14 www.beyondblue.org.au <1%
Internet Source

---

15 usenix.org <1%
Internet Source

---

16 www.mdpi.com <1%
Internet Source

---

17 www.cnn.com <1%
Internet Source

---

18 VijayaLakshmi, Balasubramanian, and Vasudev Mohan. "Kernel-based PSO and FRVM: An automatic plant leaf type detection using texture, shape, and color features", Computers and Electronics in Agriculture, 2016. <1%
Publication

---

19 "Multimodal Affective Dimension Prediction Using Deep Bidirectional Long Short-Term Memory Recurrent Neural Networks", Proceedings of the 5th International Workshop <1%

on Audio/Visual Emotion Challenge - AVEC 15, 2015.
Publication

---

20 Figueiredo, I.N., S. Kumar, C.M. Oliveira, J.D. Ramos, and B. Engquist. "Automated lesion detectors in retinal fundus images", Computers in Biology and Medicine, 2015. <1%
Publication

---

21 Vasconcelos, C., A. Félix, and T. M. Cunha. "Preoperative assessment of deep myometrial and cervical invasion in endometrial carcinoma: Comparison of magnetic resonance imaging and histopathologic evaluation", Journal of Obstetrics and Gynaecology, 2007. <1%
Publication

---

22 academic.odysci.com <1%
Internet Source

---

23 Lopez-Otero, Paula, Laura Docio-Fernandez, and Carmen Garcia-Mateo. "Assessing speaker independence on a speech-based depression level estimation system", Pattern Recognition Letters, 2015. <1%
Publication

---

24 soar.wichita.edu <1%
Internet Source

EXCLUDE QUOTES	ON	EXCLUDE MATCHES	< 9 WORDS
EXCLUDE BIBLIOGRAPHY	ON		